ENTERED   OIPE

RAW SEQUENCE LISTING   DATE: 08/12/2002
PATENT APPLICATION: US/10/098,514   TIME: 15:01:01

Input Set : A:\PTO.VSK.txt
Output Set: N:\CRF4\08122002\J098514.raw

```
   3 <110> APPLICANT: Chang, Sandra P
   4        Christopher, David A
   5        Vine, Benjamin
   6        Su, Wei-Wen
   7        Bugos, Robert
   9 <120> TITLE OF INVENTION: PLASMODIUM FALCIPARUM MEROZOITE SURFACE PROTEIN-1 MALARIA
PRODUCED IN
  10        TRANSGENIC PLANTS
  12 <130> FILE REFERENCE: A-71339/RFT/TAL/NBC
  14 <140> CURRENT APPLICATION NUMBER: US 10/098,514
C--> 15 <141> CURRENT FILING DATE: 2002-08-06
  17 <150> PRIOR APPLICATION NUMBER: US 09/500,376
  18 <151> PRIOR FILING DATE: 2000-02-08
  20 <150> PRIOR APPLICATION NUMBER: US 60/274,599
  21 <151> PRIOR FILING DATE: 2001-03-09
  23 <160> NUMBER OF SEQ ID NOS: 19
  25 <170> SOFTWARE: PatentIn version 3.1
  27 <210> SEQ ID NO: 1
  28 <211> LENGTH: 1149
  29 <212> TYPE: DNA
  30 <213> ORGANISM: Artificial sequence
  32 <220> FEATURE:
  33 <223> OTHER INFORMATION: synthetic
  35 <220> FEATURE:
  36 <221> NAME/KEY: sig_peptide
  37 <222> LOCATION: (1)..(24)
  38 <223> OTHER INFORMATION:
  41 <400> SEQUENCE: 1
  42 gccgaattcg acaacatcct cagtgacaac atcctcagtg gcttcgagaa cgagtacgac      60
  44 gtaatctacc taaagcccct tgccggtgtc taccgttcat tgaagaaaca gatagaaaag     120
  46 aatatttttca cgttcaacct caacctaaat gacatcctca actcgcgcct caagaagcga     180
  48 aaatacttcc tcgacgtgtt ggaatccgac cttatgcaat tcaagcacat tagctctaac     240
  50 gagtacatca tagaggacag cttcaagctc ttgaattcag aacagaagaa caccctccta     300
  52 aagtcctaca aatacattaa ggagtctgtt gagaacgaca tcaagttcgc ccaggaagga     360
  54 attagctact atgagaaagt cctggctaaa tacaaggacg acttggaaag cattaagaag     420
  56 gtaatcaaag aagagaagga aaagtttccg agctctccac ccacaactcc cccatcgcct     480
  58 gcaaagaccg acgagcagaa aaaagaaagt aagttccttc cattcctcac caacatcgaa     540
  60 actctatata acaacctggt gaacaagatt gatgactact aatcaactt gaaggcgaaa     600
  62 attaatgact gtaacgtcga aaaggatgaa gcccacgtta agatcaccaa gctttccgat     660
  64 ctcaaagcca tcgacgataa gattgacctg tttaagaacc acaacgattt cgacgcaatc     720
  66 aaaaagttga tcaacgacga tactaagaaa gacatgcttg gaaaactgct gtcgacaggc     780
  68 ttggtccaaa acttcccgaa caccattata agcaagctga tcgaaggaaa gtttcaggat     840
  70 atgctgaaca tctctcagca tcaatgcgtg aagaagcaat gtcccgagaa ttcaggttgc     900
  72 ttccgccact tagacgaaag ggaggaatgt aaatgcctgc tgaattataa acaggaagga     960
```

RAW SEQUENCE LISTING   DATE: 08/12/2002
PATENT APPLICATION: US/10/098,514   TIME: 15:01:01

Input Set : A:\PTO.VSK.txt
Output Set: N:\CRF4\08122002\J098514.raw

```
74 gacaagtgcg tagagaatcc taacccaacc tgtaacgaaa ataacggtgg ctgcgatgct   1020
76 gacgctaagt gtaccgagga ggacagcggt tccaatggca agaaaataac ttgcgaatgc   1080
78 acgaagcccg atagttaccc tctcttcgac ggtatcttct gctcccatga tgagctttaa   1140
80 gagctcacc                                                           1149
83 <210> SEQ ID NO: 2
84 <211> LENGTH: 383
85 <212> TYPE: PRT
86 <213> ORGANISM: Artificial sequence
88 <220> FEATURE:
89 <223> OTHER INFORMATION: synthetic
91 <220> FEATURE:
92 <221> NAME/KEY: MISC_FEATURE
93 <222> LOCATION: (380)..(380)
94 <223> OTHER INFORMATION: "Xaa" at position 380 represents a stop codon
97 <400> SEQUENCE: 2
99  Ala Glu Phe Asp Asn Ile Leu Ser Asp Asn Ile Leu Ser Gly Phe Glu
100 1               5                  10                  15
103 Asn Glu Tyr Asp Val Ile Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg
104             20                  25                  30
107 Ser Leu Lys Lys Gln Ile Glu Lys Asn Ile Phe Thr Phe Asn Leu Asn
108         35                  40                  45
111 Leu Asn Asp Ile Leu Asn Ser Arg Leu Lys Lys Arg Lys Tyr Phe Leu
112     50                  55                  60
115 Asp Val Leu Glu Ser Asp Leu Met Gln Phe Lys His Ile Ser Ser Asn
116 65                  70                  75                  80
119 Glu Tyr Ile Ile Glu Asp Ser Phe Lys Leu Leu Asn Ser Glu Gln Lys
120             85                  90                  95
123 Asn Thr Leu Leu Lys Ser Tyr Lys Tyr Ile Lys Glu Ser Val Glu Asn
124             100                 105                 110
127 Asp Ile Lys Phe Ala Gln Glu Gly Ile Ser Tyr Tyr Glu Lys Val Leu
128         115                 120                 125
131 Ala Lys Tyr Lys Asp Asp Leu Glu Ser Ile Lys Lys Val Ile Lys Glu
132     130                 135                 140
135 Glu Lys Glu Lys Phe Pro Ser Ser Pro Pro Thr Thr Pro Pro Ser Pro
136 145                 150                 155                 160
139 Ala Lys Thr Asp Glu Gln Lys Lys Glu Ser Lys Phe Leu Pro Phe Leu
140             165                 170                 175
143 Thr Asn Ile Glu Thr Leu Tyr Asn Asn Leu Val Asn Lys Ile Asp Asp
144             180                 185                 190
147 Tyr Leu Ile Asn Leu Lys Ala Lys Ile Asn Asp Cys Asn Val Glu Lys
148         195                 200                 205
151 Asp Glu Ala His Val Lys Ile Thr Lys Leu Ser Asp Leu Lys Ala Ile
152     210                 215                 220
155 Asp Asp Lys Ile Asp Leu Phe Lys Asn His Asn Asp Phe Asp Ala Ile
156 225                 230                 235                 240
159 Lys Lys Leu Ile Asn Asp Asp Thr Lys Lys Asp Met Leu Gly Lys Leu
160             245                 250                 255
163 Leu Ser Thr Gly Leu Val Gln Asn Phe Pro Asn Thr Ile Ile Ser Lys
164             260                 265                 270
```

RAW SEQUENCE LISTING   DATE: 08/12/2002
PATENT APPLICATION: US/10/098,514   TIME: 15:01:01

Input Set : A:\PTO.VSK.txt
Output Set: N:\CRF4\08122002\J098514.raw

```
     167 Leu Ile Glu Gly Lys Phe Gln Asp Met Leu Asn Ile Ser Gln His Gln
     168         275                 280                 285
     171 Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly Cys Phe Arg His Leu
     172         290                 295                 300
     175 Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly
     176 305                 310                 315                 320
     179 Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly
     180             325                 330                 335
     183 Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn
     184         340                 345                 350
     187 Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu
     188         355                 360                 365
W--> 191 Phe Asp Gly Ile Phe Cys Ser His Asp Glu Leu Xaa Glu Leu Thr
     192     370                 375                 380
```

     195 <210> SEQ ID NO: 3
     196 <211> LENGTH: 1149
     197 <212> TYPE: DNA
     198 <213> ORGANISM: Artificial sequence
     200 <220> FEATURE:
     201 <223> OTHER INFORMATION: synthetic
     203 <220> FEATURE:
     204 <221> NAME/KEY: sig_peptide
     205 <222> LOCATION: (1)..(24)
     206 <223> OTHER INFORMATION:
     209 <400> SEQUENCE: 3

```
     210 cggatccaag gagatataac aatggacaac atcctcagtg gcttcgagaa cgagtacgac      60
     212 gtaatctacc taaagcccct tgccggtgtc taccgttcat tgaagaaaca gatagaaaag     120
     214 aatattttca cgttcaacct caacctaaat gacatcctca actcgcgcct caagaagcga     180
     216 aaatacttcc tcgacgtgtt ggaatccgac cttatgcaat tcaagcacat tagctctaac     240
     218 gagtacatca tagaggacag cttcaagctc ttgaattcag aacagaagaa caccctccta     300
     220 aagtcctaca aatacattaa ggagtctgtt gagaacgaca tcaagttcgc ccaggaagga     360
     222 attagctact atgagaaagt cctggctaaa tacaaggacg acttggaaag cattaagaag     420
     224 gtaatcaaag aagagaagga aaagtttccg agctctccac ccacaactcc cccatcgcct     480
     226 gcaaagaccg acgagcagaa aaaagaaagt aagttccttc cattcctcac caacatcgaa     540
     228 actctatata acaacctggt gaacaagatt gatgactact taatcaactt gaaggcgaaa     600
     230 attaatgact gtaacgtcga aaaggatgaa gcccacgtta agatcaccaa gctttccgat     660
     232 ctcaaagcca tcgacgataa gattgacctg tttaagaacc acaacgattt cgacgcaatc     720
     234 aaaaagttga tcaacgacga tactaagaaa gacatgcttg aaaactgct gtcgacaggc      780
     236 ttggtccaaa acttccccgaa caccattata agcaagctga tcgaaggaaa gtttcaggat     840
     238 atgctgaaca tctctcagca tcaatgcgtg aagaagcaat gtcccgagaa ttcaggttgc     900
     240 ttccgccact tagacgaaag ggaggaatgt aaatgcctgc tgaattataa acaggaagga     960
     242 gacaagtgcg tagagaatcc taacccaacc tgtaacgaaa ataacggtgg ctgcgatgct    1020
     244 gacgctaagt gtaccgagga ggacagcggt tccaatggca agaaaataac ttgcgaatgc    1080
     246 acgaagcccg atagttaccc tctcttcgac ggtatcttct gctcccatga tgagctttaa    1140
     248 gagctcacc                                                          1149
```

     251 <210> SEQ ID NO: 4
     252 <211> LENGTH: 383
     253 <212> TYPE: PRT
     254 <213> ORGANISM: Artificial sequence

```
256 <220> FEATURE:
257 <223> OTHER INFORMATION: synthetic
259 <220> FEATURE:
260 <221> NAME/KEY: MISC_FEATURE
261 <222> LOCATION: (380)..(380)
262 <223> OTHER INFORMATION: "Xaa" at position 380 represents a stop codon
265 <400> SEQUENCE: 4
267 Arg Ile Gln Gly Asp Ile Thr Met Asp Asn Ile Leu Ser Gly Phe Glu
268 1               5                10                    15
271 Asn Glu Tyr Asp Val Ile Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg
272             20                  25                  30
275 Ser Leu Lys Lys Gln Ile Glu Lys Asn Ile Phe Thr Phe Asn Leu Asn
276         35                  40                  45
279 Leu Asn Asp Ile Leu Asn Ser Arg Leu Lys Lys Arg Lys Tyr Phe Leu
280     50                  55                  60
283 Asp Val Leu Glu Ser Asp Leu Met Gln Phe Lys His Ile Ser Ser Asn
284 65                  70                  75                  80
287 Glu Tyr Ile Ile Glu Asp Ser Phe Lys Leu Leu Asn Ser Glu Gln Lys
288                 85                  90                  95
291 Asn Thr Leu Leu Lys Ser Tyr Lys Tyr Ile Lys Glu Ser Val Glu Asn
292             100                 105                 110
295 Asp Ile Lys Phe Ala Gln Glu Gly Ile Ser Tyr Tyr Glu Lys Val Leu
296         115                 120                 125
299 Ala Lys Tyr Lys Asp Asp Leu Glu Ser Ile Lys Lys Val Ile Lys Glu
300     130                 135                 140
303 Glu Lys Glu Lys Phe Pro Ser Ser Pro Pro Thr Thr Pro Pro Ser Pro
304 145                 150                 155                 160
307 Ala Lys Thr Asp Glu Gln Lys Lys Glu Ser Lys Phe Leu Pro Phe Leu
308             165                 170                 175
311 Thr Asn Ile Glu Thr Leu Tyr Asn Asn Leu Val Asn Lys Ile Asp Asp
312             180                 185                 190
315 Tyr Leu Ile Asn Leu Lys Ala Lys Ile Asn Asp Cys Asn Val Glu Lys
316         195                 200                 205
319 Asp Glu Ala His Val Lys Ile Thr Lys Leu Ser Asp Leu Lys Ala Ile
320     210                 215                 220
323 Asp Asp Lys Ile Asp Leu Phe Lys Asn His Asn Asp Phe Asp Ala Ile
324 225                 230                 235                 240
327 Lys Lys Leu Ile Asn Asp Asp Thr Lys Lys Asp Met Leu Gly Lys Leu
328             245                 250                 255
331 Leu Ser Thr Gly Leu Val Gln Asn Phe Pro Asn Thr Ile Ile Ser Lys
332             260                 265                 270
335 Leu Ile Glu Gly Lys Phe Gln Asp Met Leu Asn Ile Ser Gln His Gln
336         275                 280                 285
339 Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly Cys Phe Arg His Leu
340     290                 295                 300
343 Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly
344 305                 310                 315                 320
347 Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly
348             325                 330                 335
```

```
       351 Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn
       352         340                 345                 350
       355 Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu
       356         355                 360                 365
W-->   359 Phe Asp Gly Ile Phe Cys Ser His Asp Glu Leu Xaa Glu Leu Thr
       360         370                 375                 380
       363 <210> SEQ ID NO: 5
       364 <211> LENGTH: 38
       365 <212> TYPE: DNA
       366 <213> ORGANISM: Artificial sequence
       368 <220> FEATURE:
       369 <223> OTHER INFORMATION: synthetic
       371 <400> SEQUENCE: 5
       372 cgggatcccca aggagatacc atggcagtaa ctccttcc                    38
       375 <210> SEQ ID NO: 6
       376 <211> LENGTH: 42
       377 <212> TYPE: DNA
       378 <213> ORGANISM: Artificial sequence
       380 <220> FEATURE:
       381 <223> OTHER INFORMATION: synthetic
       383 <400> SEQUENCE: 6
       384 ctgcgagctc ttattaaagc tcatcatgct gcagaaaata cc               42
       387 <210> SEQ ID NO: 7
       388 <211> LENGTH: 28
       389 <212> TYPE: DNA
       390 <213> ORGANISM: Artificial sequence
       392 <220> FEATURE:
       393 <223> OTHER INFORMATION: synthetic
       395 <400> SEQUENCE: 7
       396 ctgcgagctc ttattaatga tgatgatg                               28
       399 <210> SEQ ID NO: 8
       400 <211> LENGTH: 42
       401 <212> TYPE: DNA
       402 <213> ORGANISM: Artificial sequence
       404 <220> FEATURE:
       405 <223> OTHER INFORMATION: synthetic
       407 <400> SEQUENCE: 8
       408 ggtgagctct taaagctcat catgggagca gaagataccg tc               42
       411 <210> SEQ ID NO: 9
       412 <211> LENGTH: 24
       413 <212> TYPE: DNA
       414 <213> ORGANISM: Artificial sequence
       416 <220> FEATURE:
       417 <223> OTHER INFORMATION: synthetic
       419 <400> SEQUENCE: 9
       420 gccgaattcg acaacatcct cagt                                   24
       423 <210> SEQ ID NO: 10
       424 <211> LENGTH: 42
       425 <212> TYPE: DNA
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 08/12/2002
PATENT APPLICATION: US/10/098,514   TIME: 15:01:02

Input Set : A:\PTO.VSK.txt
Output Set: N:\CRF4\08122002\J098514.raw

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; Xaa Pos. 380
Seq#:4; Xaa Pos. 380

VERIFICATION SUMMARY   DATE: 08/12/2002
PATENT APPLICATION: US/10/098,514   TIME: 15:01:02

        Input Set : A:\PTO.VSK.txt
        Output Set: N:\CRF4\08122002\J098514.raw

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:191 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:368
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:368